

Error Theories and Bare-Difference Methodology: A Reply to Kopeikin

Scott Hill

(This is a draft. Please cite the final version which is forthcoming in the *Journal of Value Inquiry*.)

Abstract

Kopeikin (forthcoming a, forthcoming b) and Rachels' (1975) bare-difference cases elicit the intuition that killing is no different than letting die. Hill's (2018) bare-difference cases elicit the intuition that killing is worse than letting die. At least one of the intuitions must be mistaken. This calls for an error theory. Hill has an error theory for the intuition elicited by the Kopeikin/Rachels' cases. Kopeikin and Rachels have an error theory for the intuition elicited by Hill's cases. A natural thought is that we are at an impasse. There is no plausible basis for preferring one error theory to the other. I argue that this natural thought is mistaken. Not all error theories are equal. Preliminary considerations favor Hill's error theory and disfavor the Kopeikin/Rachels error theory. But preliminary considerations are not decisive. The way forward in the bare-difference debate is not to evaluate intuitions. The intuitions are in. What is left to do now is evaluate the comparative status of the Hill and the Kopeikin/Rachels error theories.

Introduction

For debunking purposes, I do not think all error theories are equal. Some are better than others. One thing an error theory can do is provide a defeater. Suppose there is some independent reason to think an intuition is untrustworthy. Or suppose that there is some independent reason to think the belief supported by an intuition is false. In such cases, support for the relevant belief is eroded. Let's say that an error theory is a *defeating error theory* for an intuition *iff* it provides a defeater for the belief that that intuition supports.

An error theory can also make salient previously unrecognized possibilities. Suppose a belief is supported by an intuition. Suppose a theory has the counterintuitive result that the relevant belief is false. But suppose also that there is a story that would explain how the relevant belief could be false. No defeater is provided. No independent reason is given for the truth of the error theory. But a previously unrecognized possibility is made salient. Let's say that an error theory is a *mere just so story* for an intuition *iff* it identifies a possibility in which the belief that that intuition supports is false.

A sufficiently strong defeating error theory warrants completely discounting the force of an intuition. This is simply for the familiar reasons that a sufficiently strong defeater warrants completely discounting the force of an intuition. A mere just so story, on the other hand, provides little reason to discount an intuition. Consider my perceptual beliefs: I see a hand in front of me. On the basis of what I see, I form the belief that there is a hand in front of me. But it is possible that I am dreaming. That would explain my perceptual experiences. But it provides me with no reason to discount the beliefs that are based on my perceptual experiences. I think that there are important roles for mere just so stories in theory construction and evaluation. And I employ them myself elsewhere. But for the purposes of debunking, my view is that a sufficiently strong defeating error theory can do the job. But a mere just so story cannot.

In this paper, I argue Hill's error theory is a defeating error theory and the Kopeikin/Rachels error theory is a mere just so story. So as things stand, bare difference methodology supports the view that killing is worse than letting die. Below, I assume familiarity with Kopeikin's coin cases, Rachels' cases, and Hill's cases. But for the reader unfamiliar with this literature, I include the relevant cases in an Appendix for reference.

Hill's Error Theory

In the coin cases, the differences between Smith's act and Jones' act are only due to a lucky difference in the outcome of coin tosses. Our intuitions about cases involving moral luck are untrustworthy. So, although I share Kopeikin's intuition about the coin cases, I think that intuition should not be trusted. In reply, Kopeikin (forthcoming a, p.) says: "the same challenge can be leveled against Hill's own cases." And I think it is plausible that if we go back far enough, luck will explain why Jones has one preference and Smith has another.

However, my view is that what leads our intuitions astray isn't that moral luck can be found somewhere eventually if we think about the case long enough. Instead, it is plausible that moral luck is present in all cases. The issue isn't whether it is objectively there. The issue is whether it is brought to the forefront and highlighted. Think about Nagel's drunk driver cases. When luck is left in the background, we think the drunk driver that hits the curb and kills a family has done something worse than the otherwise similar drunk driver that merely hits the curb. But when it is highlighted that the only difference between their acts is the result of luck, our intuition is distorted and the acts seem the same. Both drivers took the same risks, drove in exactly the same way, hit the exact same curb, and drove on the exact same sidewalk. But one of the drivers was unlucky because there happened to be a family on the sidewalk. Nagel points out that the natural reaction, after luck is highlighted is to think there really is no difference between their acts.

Consider the ethicists that are the target of Hill's earlier paper—Rachels (1986), Singer (2005), Oddie (1997), and others. They will agree that there is a difference between the two drunk drivers' acts. For the relevant acts have different consequences. And though the differences in consequences are due to luck, they are still morally relevant for the sort of ethicists that tend to be sympathetic to Rachels' argument. So even the main figures that think killing is no worse than letting die should agree with that moral luck serves to make acts that are different seem equivalent. And it isn't the objective presence of moral luck that distorts our intuitions. It is taking moral luck out of the background and making it especially salient that distorts our intuitions.

Now return to the coin cases. Constructing the cases in such a way that the difference between acts is due entirely to the outcome of a coin toss puts moral luck front and center and makes it maximally salient. Constructing the cases in such a way that the differences in acts are due to differences in preferences, as in Hill's examples, puts moral luck in the background. Our intuitions are reliable when moral luck is in the background but unreliable when moral luck is salient. So we should think our intuitions about Hill's cases are more reliable than our intuitions about the coin cases.

In light of this, Hill's error theory is independently motivated. And it is an error theory that the utilitarians that find Rachels cases compelling are forced to accept. So Hill's error theory is a defeat error theory.

There is a further point to consider. When moral luck is made salient, it doesn't mislead our intuitions in a random way. There is a specific way in which it is misleading. In particular, it mutes the difference between acts. So to the extent to which moral luck is salient in Hill's cases, we should think that it is muting the difference between killing and letting die. My considered view, then, is this: Grant for the sake of argument that moral luck is at least a little bit salient in Hill's cases. We have the intuition that killing is worse than letting die. Moral luck mutes the difference between acts. So the presence of moral luck in Hill's cases should make us think that we are *underestimating* the difference between killing and letting die. On the other hand, moral luck is maximally salient in the coin cases. And we lack the intuition that killing is worse than letting die in those cases. But that provides no evidence for the view that killing is equivalent to letting die. For even if the relevant acts are different, the maximal salience of moral luck would make it seem that they are the same. In the end, I think that if we grant that moral luck is at least a bit salient in Hill's cases, then that only shows that we should be even more confident in the intuition that killing is worse than letting die. And I think that is consistent with holding that moral luck should cause us to disregard our intuitions about the coin cases.

Kopeikin (forthcoming a, p.) makes another relevant point in connection with moral luck: "it's not clear that the coin-flip does introduce moral luck. It strikes me that one would invoke moral

luck in my cases only if one antecedently thought killing is worse than letting die.” So even if I am right that Hill’s error theory defeats intuitions generated by cases in which moral luck is salient, it might be that the error theory is inapplicable to the coin cases.

I agree that moral luck is present in the coin cases only if there is a difference between killing and letting die. But I think that is sufficient to support the use of Hill’s cases rather than the coin cases. For whether or not killing is worse than letting die, the coin cases will give the same result. If killing and letting die are equivalent, then there is no moral luck, and we’ll have the intuition that they are equivalent. If killing is worse than letting die, then there is moral luck, and our intuitions will be distorted in such a way that the acts seem equivalent. For this reason, it is not a helpful test. It is insensitive to what we are testing for. Hill’s cases do not make moral luck salient. And so they are not insensitive to whether killing is worse than letting die in the way that the coin cases are.

The Kopeikin/Rachels Error Theory

Smith prefers to kill. Jones prefers to let die. Hill thinks, and Kopeikin (forthcoming a, p.) grants for the sake of argument, that Smith’s preference is worse than Jones’. Kopeikin rightly worries that Hill might be confusing the comparative badness of Smith and Jones’ acts with the comparative badness of their preferences. However, Hill has a natural explanation of why Smith’s preference is worse than Jones’. It is that Smith prefers to perform an act—killing—that is worse than the act Jones prefers to perform—letting die. The opponent of the view that killing is worse than letting die cannot offer this explanation. And I don’t see what explanation they could give. So I don’t see that the fact that Smith’s preference is worse provides a basis for thinking that Hill is confusing act with preference. We need to explain why Smith’s preference is worse. And I don’t see how we can do that without supposing that killing is worse than letting die.

Kopeikin’s (forthcoming a, p.) strategy is to deny the claim that Smith’s preference is worse than Jones’. Here is where the Kopeikin/Rachels error theory comes in. As Kopeikin (forthcoming, p.) puts it:

Rachels writes that we regularly hear from the media of horrible cases of killing, but rarely hear of letting-die cases and that, when we do, they involve doctors ‘motivated by humanitarian reasons’ [1975: 80]. Given this kind of exposure, one plausibly ‘learns to think of killing in a much worse light than of letting die’ [ibid.]. I think that Rachels is correct, but more needs to be said.

And then Kopeikin goes on to fill in the details about how this might work. This error theory explains why we think killing is worse than letting die. And it is the misguided intuition that killing is worse than letting die that makes us think Smith’s preference is worse. So an explanation for why Smith’s preference is worse is no longer needed.

I think Kopeikin’s error theory is very interesting and worthy of further development. However, I also think the Kopeikin/Rachels error theory, in its current stage of development, is not sufficiently supported and is a mere just so story. It is possible that the media highlighting killing rather than letting die explains why we think killing is worse. But it is also possible that the media highlights killing rather than letting die because they know we think killing is far worse. And they know highlighting the far worse thing will get attention and get clicks.

I have argued that the Hill error theory is a defeating error theory that is sufficiently strong to completely undermine our intuitions in the coin cases and in Rachels’ original cases. In contrast, the Kopeikin/Rachels error theory, as best as I can tell, just highlights a possibility. Kopeikin is in good company. I think it is common for applied ethicists to appeal to mere just so stories. But I also think that such stories, by themselves, provide little reason to discount our intuitions.

Character and Act

A common move in the literature on this topic is to hold that insofar as we have the intuition that killing and letting die are different in a pair of cases, we are confusing character and act. As we have

seen, Kopeikin (forthcoming a) and (forthcoming b) makes this move. And he is in good company. Kuhse (1998) and Perrett (1996) make this move too in especially important papers on this topic.

Suppose we agree with the widespread thought that it is tempting to confuse character and act evaluation. If that is a possible worry about Hill's cases, it is also a possible worry about the Kopeikin/Rachels cases. In their cases, Jones and Smith have exactly the same desires and dispositions to act. So their characters are the same. Just as differences in character might tempt us to find differences in acts when those differences are not there, similarity in character might tempt us to think that acts are morally similar when they are really different. So the allegation of confusing character with act is a double-edged sword. Though most popularly wielded against the view that killing is worse than letting die, it is at least as potent against the view that killing and letting die are equivalent.

Furthermore, I have offered a response to the worry that Hill is confusing act with character. Even if he is, the best explanation of why characters are different in Hill's cases is that acts are different in Hill's cases. But that same explanation is not available to Kopeikin and Rachels.

This raises the question of whether there is something off about bare-difference methodology in the first place. Bare difference cases are supposed to be ones in which everything is the same across cases with the exception that one is a case of killing the other is a case of letting die. But if everyone agrees that in all of the examples in the literature the characters are different, doesn't that imply that it is impossible to construct bare difference cases?

I don't think that is quite right. I think this shows that Rachels' original statement of what bare difference cases are wasn't refined enough. I think there are at least two ways of refining Rachels' account of bare difference cases that are worth considering and that do not have this problem.

BDC₁: Two cases are bare difference cases *iff* they are as similar as is consistent with one being an act of killing and the other being an act of letting die.

BDC₂: Two cases are bare difference cases *iff* (i) they are identical with the exception that one is a case of killing and the other is a case of letting die and (ii) any other differences in the cases are due to that one difference.

If I am right about what is going on in Hill's cases, they are bare difference cases according to either account just given. More work needs to be done to show that these can do the work that the original account of bare difference cases is supposed to do. But this is a first pass at how to address the relevant issue. Furthermore, if it turns out that it is impossible to construct bare-difference cases, then I think that is a win for the proponent of the view that killing is worse than letting die. The point of bare-difference methodology was originally to refute the view that killing is worse than letting die. If bare different cases cannot be constructed, then that refutation fails.

Violence

Hill's cases differ with respect to violence. In one case there is an extra throat slitting. In the other case there is not. Kopeikin (forthcoming a, p.) worries that this asymmetry might mislead our intuitions. I think he is absolutely right to point that out. However, there are examples that elicit the same intuition as Hill's cases but do not differ with respect to violence. Consider *Oddie's Trolley*:

"The conclusion of the argument has an obvious but possibly disturbing implication. A runaway trolley with one passenger in it is careering out of control down a track towards the left-hand fork further along which sits the stationary but this time empty trolley. A collision will cause the death of the passenger on the moving trolley. The right-hand fork still leads to the siding with the brick wall at the end of it. You could either do nothing, thereby letting one person die, or you could send the trolley into the brick wall, thereby intervening to cause

his death. You could either let him die or you could kill him. According to equivalence, there is no value-difference between the two.”

Oddie’s (1997, p. 273) example is better than Hill’s in the respect Kopeikin identifies. The two alternatives—kill or let die—do not differ with respect to their degree of violence. Each is death by a violent crash. Oddie defends the view that killing and letting die are equivalent. But he worries that his example poses a problem for his view. And, as Hill (2018) discusses, Oddie attempts to “explain away the recalcitrant intuition” that killing is worse in Oddie’s case. Moreover, consider a variant of Hill’s examples that removes violence altogether:

Accident (Gentle Gas): Jones will gain a large inheritance if his six-year old cousin dies. Jones decides to murder his cousin by releasing a gentle but deadly gas while the child is bathing. But, as Jones begins to release the gas, he sees that an independently caused gas leak will kill his cousin. Jones stands ready to release the gas if necessary. But the child dies as a result of the independent leak. And Jones does nothing.

And

Accident + Murder (Gentle Gas): As before, Jones gets ready to release the gas and kill his cousin and, before he does so, he notices an independent leak which, if nothing is done, will kill his cousin. But Jones refuses to let his cousin die. He insists on patching up the leak before it can do any harm and then releasing his own gas and killing the child himself.

It seems to me that if one agreed that killing is worse than letting die in Hill’s original cases, then one should agree that killing is worse than letting die in the *Gentle Gas* variants of Hill’s cases. Both acts are bad. But killing, in the *Gentle Gas* variants, is at least a bit worse.

Finally, let me a new pair of examples:

Allowing: Over the course of four years you spend \$3.4K on video games, fancy beer, dental work, internet service, and books for your children by pressing buttons when you could have pressed other buttons and spent that money to save a child’s life.

Doing: Over the course of four years you get \$3.4K in video games, fancy beer, dental work, internet service, and books for your children in exchange for pressing some buttons that cause a child to die.

In each case you act by pressing buttons. But in one case you kill and in the other case you allow death. It seems to me that the former is permissible and the latter is impermissible. There is no difference with respect to violence in these cases as in Hill’s original cases. Moral luck is not made salient as in the Kopeikin/Rachels cases. I do not think we should allow our intuitions about cases like *Allowing* and *Doing* to be overturned by appeal to examples that highlight moral luck.

Objections and Replies

First Objection: In the drunk driver cases, there is a difference in consequences. In the coin cases, there is no difference in consequences.

Reply: I don’t see how this asymmetry is a problem for my argument. Suppose the acts in a pair of cases differ due to moral luck. Then, recall, that my argument depends on the following principle:

Luck Salience: As moral luck is made more and more salient, the acts seem more and more equivalent.

And this is true whatever differences there are between the acts. If the consequences, character, whatever are different, as moral luck is made more and more salient, the acts seem the same. If the only difference is that one is a case of doing and the other is a case of allowing, then as moral luck is made more and more salient, the acts seem more and more the same. And so, I claim, we can't use cases in which moral luck is made especially salient to test whether acts are morally different. For whether they really are morally different or not, highlighting moral luck will flatten that difference out. Making the difference between acts due to a coin flip makes the moral luck maximally salient and so maximally flattens the seeming difference between the acts even if they are different.

Second Objection: I (the objector) lack the intuition that the drunk driver cases are the same. The acts seem different even after the fact that they are due to luck is made salient.

Reply: Fair enough. But there is a large body of literature developed around intuitions that others have that making moral luck salient does mute the intuitive difference between two otherwise different acts. And this is also present in discussions of the ethics of belief. Consider Clifford's boat examples. Two captains who suppress their doubts about the seaworthiness of their ships and send their ships out. One is lucky and his ship sails safely. The other is unlucky and his ship sinks. Clifford presses this point about luck into service of the claim that their acts of belief are equally bad. So although some particular philosophers may lack the intuition that the drunk drivers or that Clifford's captains acts and beliefs are equivalent after moral luck is made salient, that does not remove the fact that many philosophers report that their intuitions that the acts are different is substantially muted and even eliminated in such cases. And if we agree that the relevant acts are actually different, or at least that our intuitions in such cases jump back and forth between there being a difference and there not being a difference, then we should be skeptical of examples that put moral luck so far at the forefront as the coin examples do.

Third Objection: I (the objector) agree that the act in *Accident + Murder (Gentle Gas)* is worse than the act in *Accident (Gentle Gas)*. However, there is a problem with the killing variation of Hill's case. By taking the time and effort to patch up the leak so that he can kill his cousin, it seems as though Jones is more motivated to be the hand by which his cousin dies than in the first case, where he does nothing. And that is why the act in *Accident + Murder (Gentle Gas)* is worse.

Reply: Hill addresses a similar worry in reply to an objection concerning his original examples. And I address a similar worry earlier in this paper. The core idea is this: Suppose being more motivated to bring about the death of Jones' cousin by one's own hand makes the act of killing worse than the act of letting die. Then we need an explanation of why this is worse. The proponent of the view that killing is worse than letting die has such an explanation. In particular, being more motivated to bring about death by one's own hand is worse because killing is worse. If killing isn't any worse than letting die, then it is puzzling why the increased desire to kill would be worse. Compare: There is no difference between killing in a green t-shirt and killing in a red t-shirt. And if I have the choice to kill in either shirt but insist on the green shirt, that doesn't make my act any worse. Furthermore, there is no difference between effort spent in the two cases. In *Accident (Gentle Gas)*, Jones takes extra effort and time to patch up his own gas leak so that he isn't the hand by which his cousin dies. The difference between the cases is that in one Jones makes extra effort to avoid killing his cousin and merely let him die. In the other case, Jones makes extra effort to be the hand by which his cousin dies.

Furthermore, whatever complications may arise due to this factor in the Gentle Gas variants of Hill's original examples, no more effort is made in *Allowing* than *Doing*. But our intuition is that the acts are different.

Fourth Objection: I claim that *Luck Salience* is true and that making moral luck salient distorts our intuitions. I claim that it takes acts that are different and makes them seem the same. And it is true, as an empirical matter, that making moral luck salient *changes* our intuitions. However, there is an alternative moral hypothesis that accommodates the empirical data but does not entail that making moral luck salient misleads us. Moral luck cases shift our focus from evaluations of act to evaluations of character. Our initial intuitions about moral luck cases, such as the drunk driver cases, are that the acts are different, but once the moral luck is made more salient, our intuitions shift to recognizing

that the character of the agents is equivalent (because, for example, both people took unacceptable risks of driving drunk) and this explains how moral luck affects our intuitions.

Reply: I do not think of this hypothesis as an alternative to *Luck Salience*. Rather, I think the objector provides a more precise filling out of *Luck Salience* than what I offer above and a nice explanation of why *Luck Salience* is true. And I think the objector's point lends further support to my claim that Hill's examples are to be preferred in bare difference reasoning and the Kopeikin/Rachels cases are to be avoided.

Remember, the objector's idea here is that when moral luck is in the background, we are eliciting intuitions about acts. And when moral luck is brought front and center, we are eliciting intuitions about character. If the objector is right, then Kopeikin's coin cases and Rachels' bathtub cases, since they put moral luck front and center, elicit intuitions about character. And for that reason, if we want to know whether the act of killing is worse than the act of letting die, we shouldn't look to their cases since they elicit intuitions about character evaluation rather than act evaluation. But Hill's cases on the other hand, put moral luck in the background. And so, given the objector's hypothesis, they elicit intuitions about acts. Since we are evaluating acts and not character, it seems to me that we should use Hill's cases and not the Kopeikin/Rachels' cases.

Appendix: Relevant Cases from the Bare-Difference Literature

Rachels' Cases

Murder: Jones will gain a large inheritance if his six-year-old cousin dies. One evening, while the child is taking a bath, Jones sneaks into the bathroom, drowns the child, and makes it look like an accident.

Accident: As before, Jones sneaks into the bathroom planning to drown the child. But, as Jones enters, the child hits his head and falls face down into the water. Jones stands ready to kill the child if necessary. But the child dies on his own.

Hill's Case

Accident + Murder: As before, Jones sneaks into the bathroom planning to drown his cousin, and, as Jones enters, the child hits his head and falls face down in the water. But Jones refuses to let his cousin die. Instead, Jones insists on killing him.

Kopeikin's Cases

Heads-Indifference: Smith sneaks in and sees his cousin lying unconscious in the water drowning. Smith stops. He is indifferent to whether or not his cousin dies by his own hand. Smith removes a coin from his pocket and decides to flip it. If 'heads' is the result, he'll drown the cousin; if 'tails', he'll let the cousin drown. The coin lands as 'heads'. Smith shrugs and proceeds with the outcome of the coin-flip. He gently holds down the unconscious child until the child dies.

Tails-Indifference: Up to the coin-flip, everything proceeds in the same way for Jones. His coin lands 'tails'. He shrugs and proceeds with the outcome of the coin-flip: he stands by, ready to gently hold down the unconscious child, but doesn't need to do so. The child drowns as Jones watches.

Acknowledgments

For comments and discussion I am very grateful to Bruce Blackshaw, Zak Kopeiken, Joonas Rasanen and the referees for AJP and JVI.

References

- Hill, S. 2018. Murdering and Accident Victim: A New Objection to the Bare-Difference Argument *Australasian Journal of Philosophy* 96: 767-778
- Kopeikin, Z. forthcoming a. *Australasian Journal of Philosophy* 1-10
- Kopeikin, Z. forthcoming b. Bare-Difference Methodology and a Problematic Separability Principle *Journal of Value Inquiry* 1-18
- Kopeikin, Z. 2020. A Separability Principle, Contrast Cases, and Contributory Dispositions *Southwest Philosophy Review* 36:35-44
- Oddie, G. and P. Menzies 1992. An Objectivist's Guide to Subjective Value, *Ethics* 102: 512–33.
- Nesbitt, W. 1995. Is Killing No Worse Than Letting Die? *Journal of Applied Philosophy* 12: 101–6.
- Oddie, G. 1997. Killing and Letting-Die: Bare Differences and Clear Differences, *Philosophical Studies* 88: 267–87.
- Perrett, R. 1996. Killing, Letting Die, and the Bare Difference Argument *Bioethics* 10: 131–9.
- Rachels, J. 1986. *The End of Life: Euthanasia and Morality*, Oxford: Oxford University Press.
- Singer, P. 2005. Ethics and Intuitions *Journal of Ethics* 9: 331-352